Our Ref.: 002717.P029C

APPLICATION FOR UNITED STATES LETTERS PATENT

FOR

Method and System for VMAN Protocol Layer-2 Packet Nested Encapsulation

Inventor(s): **Michael Yip**
Steve Haddock

Prepared by:

BLAKELY SOKOLOFF TAYLOR & ZAFMAN, LLP
12400 Wilshire Boulevard, 7th Floor
Los Angeles, California 90025
(503) 684-6200

"Express Mail" label number EV325527312US

Method and System for VMAN Protocol

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to the field of virtual metropolitan area network (VMAN) topologies and internetwork communications technologies. In particular, the present invention relates to a protocol for use in a VMAN network architecture to route and forward data packets according to the VMAN configuration.

2. Background Information and Description of Related Art

A metropolitan area network (MAN) is a network that interconnects users with computer resources in a geographic area or region larger than that covered by a local area network but smaller than the area covered by a wide area network. The term is typically applied to the interconnection of networks in a city into a single larger network (which may then also offer efficient connection to a wide area network). The amount of data traffic being sent over MANs is increasing at an exponential rate. This is due in part to the increasingly ubiquitous use of the Internet by consumers and businesses, as well as the increasingly bandwidth-intensive nature of the Internet applications that are being deployed.

An important aspect of MANs is the ability of MAN service providers to create virtual private network network connections (VPNs) across a single MAN infrastructure, referred to as a virtual metropolitan area network (VMAN). VMANs allow customers having multiple locations within a metropolitan area to transport

private traffic, including virtual local area network (VLAN) traffic, over the shared single MAN.

However, the use of vMANs to handle traffic from multiple customers over a single MAN creates access and security issues. Therefore, it is important to segregate one customer from another so that there is no co-mingling of traffic.

In addition, customer traffic must be transported over the MAN without interfering with the customers' own higher-layer protocols such as DECnet, or private IP subnets. For example, the DECnet Phase IV protocol can cause problems when routed to a Layer 2 MAN because the DECnet protocol changes the media access control (MAC) address in the packet's datalink header. Since duplicate MAC addresses are typically not allowed, MAN service providers end up managing DECnet streams by hand - something which neither the provider nor the customer wants.

Accordingly, a new approach is needed to securely manage traffic in a VMAN network architecture while not interfering with higher level protocols.

SUMMARY

According to one aspect of the invention, a method and system is provided in which a VMAN protocol is used to segregate MAN traffic at a customer and a provider domain level. A switch at the edge of the MAN encapsulates a customer data packet from an initiating 802.1Q customer domain in a new Ethernet header, which is used to specify the IEEE 802.1Q VLAN tags as determined by the customer-related VLAN configurations. A switch at the core of the MAN encapsulates the data packet further in another new Ethernet header, which is used to specify new VMAN tags as determined by the MAN service provider VMAN configurations. The nested encapsulation is repeated as necessary until the data packet is eventually forwarded to a remote switch at the edge of the MAN in accordance with the VMAN configuration, or the source and destination address in the original data packet . The remote switch strips the VMAN tags from the data packet, and forwards the stripped data packet to the receiving 802.1Q customer domain as specified in the IEEE 802.1Q VLAN tag.

According to one aspect of the invention, apparatus are provided to carry out the above and other methods.

BRIEF DESCRIPTION OF DRAWINGS

The present invention will be described by way of exemplary embodiments, but not limitations, illustrated in the accompanying drawings in which like references denote similar elements, and in which:

Figure 1 illustrates a diagram overview of a Metropolitan Area Network (MAN) configuration and corresponding data packet tagged formats in accordance with one embodiment of the present invention;

Figure 2 illustrates a more detailed diagram of the data packet tagged formats in accordance with one embodiment of the present invention;

Figure 3 illustrates a flow diagram of the operation of one embodiment of a MAN using a VMAN protocol in accordance with one embodiment of the present invention; and

Figure 4 illustrates an example implementation of a MAN using a VMAN protocol in accordance with one embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

In the following description various aspects of the present invention, a VMAN protocol method and system, will be described. Specific details will be set forth in order to provide a thorough understanding of the present invention. However, it will be apparent to those skilled in the art that the present invention may be practiced with only some or all of the described aspects of the present invention, and with or without some or all of the specific details. In some instances, well known architectures, steps, and techniques have not been shown to avoid unnecessarily obscuring the present invention. For example, specific details are not provided as to whether the

method and system is implemented in a router, server or gateway, as a software routine, hardware circuit, firmware, or a combination thereof.

Various operations will be described as multiple discrete steps performed in turn in a manner that is most helpful in understanding the present invention. However, the order of description should not be construed as to imply that these operations are necessarily performed in the order they are presented, or even order dependent. Lastly, repeated usage of the phrase "in one embodiment" does not necessarily refer to the same embodiment, although it may.

Referring now to **Figure 1**, wherein a block diagram overview of a Metropolitan Area Network (MAN) configuration and corresponding data packet tagged formats in accordance with one embodiment of the present invention is shown. As illustrated, a MAN **100** includes a first edge switch 1 **115** connects a first 802.1Q Domain-A **110** to the VMAN CORE **120**, which in turn is connected via a second edge switch 2 **125** to a second 802.1Q Domain-A **130**. In one embodiment, each of the 802.1Q Domain-As **110/130** may include one or more virtual local area networks (VLANs) belonging to a particular customer of the MAN **100**.

Data packets originating within the VLANs of the 802.1Q domains in the illustrated embodiment carry a VLAN identification that can be used to provide intra-VLAN communication using existing layer-2 VLAN forwarding mechanisms. While no other information is needed, additional constraints and layer-2 filtering can be defined as needed to improve performance and security. As there is no single official standard protocol for communication of VLAN information, existing VLAN forwarding mechanisms may be employed, including proprietary VLAN communication protocols. The method commonly used is known as "frame-tagging." In frame-tagging, packets originating from a host belonging to a VLAN acquire a VLAN identifier (VLAN ID) as they are switched onto a shared backbone

network. The VLAN ID is what enables the receiving switches to forward the packets intelligently by switching the packets to only those switches that are members of the same VLAN. A non-proprietary VLAN communication protocol has been incorporated into the Institute for Electrical and Electronics Engineers (IEEE) 802.1Q standard, whereby the VLAN ID is part of the IEEE 802.1Q header inserted between the DataLink header (i.e. the Media Access Control (MAC) header) and the frame's user data. This is also referred to as an 802.1Q tagged frame.

When the 802.1Q tagged frame is transported over the MAN **100** the MAN service provider must insure that the data packet is segregated from other customer traffic so as not to compromise the security of the packet or cause conflict with other customer traffic. In the illustrated embodiment, the packets are segregated using VMANs. A VMAN allows certain traffic designated with a VMAN identifier to be forwarded along a particular path to only destinations belonging to the same VMAN. In the illustrated embodiment, an 802.1Q tagged frame **135** represents a data packet sent by a customer having a .1Q tag with VLAN ID = 10 in the first 802.1Q Domain-A **110**. Upon receipt by the first edge switch 1 **115**, the 802.1Q tagged frame **135** is converted to a VMAN tagged frame **145** by encapsulating it in a new Ethernet data packet containing a VMAN tag with a VMAN ID = 60. The VMAN tagged frame **145** is then received by the VMAN CORE **120**, which forwards the VMAN tagged frame **145** to a second edge switch **125** according to VMAN tag with the VMAN ID = 60. Depending on the architecture of the MAN, the VMAN CORE **120** may include one or more core switches that perform routing and forwarding of network traffic based on the VMAN ID and the destination media access control (MAC) address. The encapsulation of the 802.1Q tagged frame **135** may be repeated as necessary, depending upon the architecture of the MAN. In one embodiment, the data packet sent by the customer may be an untagged frame,

such as an Ethernet frame in accordance with the IEEE 802.1D protocol. Upon receipt by the second edge switch 2 **130** the VMAN tagged frame **145** is stripped of the VMAN tag with VMAN ID = 60. The stripped VMAN tagged frame **140** is then forwarded to the proper customer in VLAN 10 according the .1Q tag VLAN ID = 10. In the case of untagged frames, the proper customer is determined from a value in the edge switch 2 **130** and the destination MAC address.

A similar process is illustrated in the 802.1Q tagged frame **155** with a .1Q tag of VLAN ID = 20. The first edge switch 1 **115** converts the 802.1Q tagged frame **155** to VMAN tagged frame **150** by adding the VMAN tag with a VMAN ID = 50 after the DataLink Header portion of the 802.1Q tagged frame **155**. Upon receipt, the second edge switch 2 **130** then converts the VMAN tagged frame **150** by stripping it of the VMAN tag with VMAN ID = 50, resulting in the 802.1Q tagged frame **160** with a .1Q tag VLAN ID = 20. The stripped VMAN tagged frame **160** is then routed to the proper customer in VLAN 20 according the .1Q tag VLAN ID = 20.

As can be seen from the foregoing description, the illustrated embodiment of the present invention makes it possible to create separate customer and provider domains for data packets transported over a single MAN **100**. Customer domain encapsulation is performed at the edge of the MAN **100** using 802.1Q frame-tagging where the VMAN-enabled switches enforce other customer-related priorities and policies, while the provider encapsulation is performed at the core of the MAN **100**. This nested encapsulation gives the MAN service provider an opportunity to create VMANs for their own use, without interfering with the flow of customer traffic.

For example, provider VMANs can be used to connect customers to third-party services such as Internet Service Providers (ISPs) or Application Service Providers (ASPs). Since the connections to the ISPs and ASPs are transported over a VMAN, the customers can easily switch ISPs without disrupting their service.

MAN service providers can also use VMANs to consolidate traffic and centralize value-added services like VPNs or managed firewalls. Rather than being forced to install and maintain equipment on or near the customer premises – an expensive, labor-intensive task – providers can aggregate customer traffic for value-added services at a central office. This not only saves on the cost of providing administrative staff, but achieves better economies of scale and gives customers more reliable service. Better economies of scale are achieved due in part to the fact that each VMAN switch can set up and manage up to 4,096 VMANs.

In the illustrated embodiment, 802.1Q tagged frames are used to encapsulate data packets at the customer domain level. However, it should be understood that other variations for determining the customer domain using IP subnets or other high-level protocols may be employed without departing from the principles of or exceeding the scope of the present invention. Moreover, while the description of the embodiments of the present invention address the method and system as it applies to use by a MAN service provider, it is appreciated by those of ordinary skill in the art that method is generally applicable to any network service provider that services multiple customers over any Internetworking application including, Local Area Networks (LANs), and Wide Area Networks (WANs).

Referring now to **Figures 2a-2c**, wherein a more detailed diagram of the VMAN protocol is illustrated in accordance with one embodiment of the present invention. **Fig. 2a** illustrates an Ethernet packet as one type of data packet that may be transported over the MAN **100**. The Ethernet packet format is known in the art and has been incorporated into the IEEE 802.3 standard for LANs using Ethernet technology. As shown, the Ethernet data packet may be up to 1514 bytes in length and comprises several well-known fields including the data link layer fields of the Media Access Control (MAC) destination address **201**, the address of the host to

which the packet is addressed and the MAC source address **202**, the address of the host from which the packet originated. The type/length field **203** indicates the type of data packet and the length of the data packet. Lastly, the user data **204** is the data field which contains the actual user data which the packet is carrying from the source to the destination.

Fig. 2b illustrates an embodiment of the 802.1Q tagged format of the previously described 802.1Q tagged frames **135**, **140**, **155**, and **160**, in further detail. The 802.1Q tagged format is known in the art and has been incorporated into the IEEE 802.1Q standard for VLAN communication. In addition to the MAC destination address **201**, MAC source address **202**, type/length **203**, and user data **204** as already described, the 802.1Q tagged frame format includes an additional 4 bytes of data in type data field **205** and .1Q tag VLAN ID **206**. In one embodiment the type field **205** is a 2-byte hexadecimal number, e.g. "H8100," that defines the type of tag being used in the data packet, in this case the .1Q tag VLAN ID **206**. This type field **205** provides the information to the switch that is necessary to interpret the contents of the .1Q tag VLAN ID **206**. In the illustrated embodiment, the .1Q tag VLAN ID **206** is also a 2-byte field, and may also be formatted as a hexadecimal number that identifies the VLAN to which the data packet belongs.

Fig. 2c illustrates an embodiment of the VMAN tagged format of the previously described VMAN tagged frames **145** and **150**, in further detail. In addition to the MAC destination address **201**, MAC source address **202**, type/length **203**, user data **204**, type **205**, and .1Q tag VLAN ID **206** as already described, the VMAN tagged frame format includes an additional 4 bytes of data in type data field **207** and VMAN tag VMAN ID **208**. In one embodiment the type field **205** is a 2-byte hexadecimal number, e.g. "H8181," that defines the type of tag being used in the data packet, in this case the VMAN tag VMAN ID **208**. The type field **207** provides

the identifying information to the switch that is necessary to interpret the contents of the VMAN tag VMAN ID **208**. In the illustrated embodiment, the VMAN tag VMAN ID **208** is also a 2-byte field, and may also be formatted as a hexadecimal number that identifies the VMAN to which the data packet belongs.

In one embodiment, the use of new Ethernet headers in the form of the type **205/207** and tag **206/208** fields constitute encapsulating the data packet into a customer domain and provider domain respectively. The encapsulation allows the MAN service providers to transport data packets over a single MAN without co-mingling different customers' traffic. In addition, the encapsulation allows the MAN service providers to aggregate traffic according to provider-defined domains.

Referring now to **Figure 3**, wherein a flow diagram of the operation of a MAN using a VMAN protocol is illustrated in accordance with one embodiment of the present invention. As shown, in process block **310** a VMAN-enabled switch located at the edge of a MAN receives an input frame (e.g. an 802.1Q tagged or 802.1D untagged frame) specifying the data packet as belonging to a particular VLAN from a particular customer domain. In process block **320**, the VMAN-enabled edge switch adds a VMAN tag to the input frame to create a new VMAN tagged frame that encapsulates the original input frame. The VMAN-enabled edge switch may incorporate logic to insure that the addition of the VMAN tag will not exceed the VMAN protocol legal frame limit. In one embodiment, the legal frame limit is 1514 bytes plus 4 additional bytes for each tag for a total of 1522 bytes. However, the method and system of the present invention does not limit the number of nested encapsulations that can be used. As a result, the frame length is the original frame length plus 4 additional bytes for each layer of encapsulation. At process block **330** the VMAN tagged frame is forwarded by a core switch of the MAN to the appropriate VMAN-enabled edge switch according the VMAN tag in the VMAN tagged frame. At

process block **340**, upon receipt of the VMAN tagged frame, the VMAN-enabled edge switch strips the VMAN tag from the VMAN tagged frame as it is no longer needed. After stripping, the remaining data packet is the input frame that was originally sent by the customer domain. At process block **350**, the receiving remote VMAN-enabled edge switch forwards the stripped frame (i.e. the original input frame) to the proper remote customer domain in accordance with the remaining.1Q tag VLAN ID in the case of an 802.1Q input frame. In the case of an untagged input frame, the receiving remote VMAN-enabled edge switch forwards the stripped frame in accordance with the destination MAC address and forwarding data stored internally to the switch.

Referring now to **Figure 4**, wherein an example implementation of a MAN using a VMAN protocol is illustrated in accordance with one embodiment of the present invention. As shown, there are two different customers, Intel and Extreme Networks, each with two different customer sites. The customer Extreme Networks is illustrated as having 802.1Q domain A with two different sites. The first Extreme Networks site **410** is located in Sunnyvale and there are two VLANs associated with the site, VLAN 1234 and VLAN 2345. The second Extreme Networks site **430** is located in Santa Clara with the same VLANs 1234, and 2345. The customer Intel is illustrated as having 802.1Q domain B with two different sites as well. The first Intel site **420** is located in Sunnyvale, and there are two VLANs associated with the site, VLAN 85, and VLAN 1234. The second Intel site **440** is located in Palo Alto, and there are the same two VLANs associated with the site, VLANs 85 and 1234. Note that both Intel and Extreme Networks customers use the same VLAN ID of 1234 to designate one of their multi-site VLANs. To avoid conflict, the VMAN Switch 1 **450** encapsulates the data packets sent from Intel Palo Alto **440** in a new packet having a VMAN tag with a VMAN ID of 889. The tagged VMAN packet (also referred to as

a frame) is received by VMAN Switch 2 **460** where the VMAN ID of 889 is stripped from the data packet, and forwarded to the Intel Sunnyvale site's **420** corresponding VLAN as specified in the VLAN ID of the underlying 802.1Q tagged packet (frame). Similarly, the VMAN Switch 2 **460** encapsulates the data packets sent from Extreme Networks Sunnyvale **410** in a new packet having a VMAN tag with a VMAN ID of 888. The tagged VMAN packet (also referred to as a frame) is received by VMAN Switch 1 **450** where the VMAN ID of 888 is stripped from the data packet, and forwarded to the Extreme Network's Santa Clara site's **430** corresponding VLAN as specified in the VLAN ID of the underlying 802.1Q tagged packet (frame). A detailed example of the content of the tagged VMAN protocol data packets for one embodiment of the present invention is shown in **Figure 4**. With reference to both **Figure 4** and **Figure 2c**, for VMAN 888, after the MAC destination and source address, the tagged VMAN packet contains a VMAN type **207** of "8181" followed by a VMAN tag **208** with VMAN ID = "888" followed by a VLAN type **205** of "8100" followed by a .1Q tag with VLAN ID = "1234" followed by the packet's type/length **203** and user data **204**. For VMAN 889, after the MAC destination and source address, the tagged VMAN packet contains a VMAN type **207** of "8181" followed by a VMAN tag **208** with VMAN ID = "889" followed by a VLAN type **205** of "8100" followed by a .1Q tag with VLAN ID = "1234" followed by the packet's type/length **203** and user data **204**. As is shown, the data packets are securely transported to the proper destination as a result of the VMAN protocol processing which segregates the Intel from the Extreme Networks traffic by using the different VMAN designations.

Accordingly, a novel method and system is described for a VMAN protocol used in forwarding data packets by a MAN switch connecting multiple customers across a single MAN infrastructure. From the foregoing description, those skilled in

the art will recognize that many other variations of the present invention are possible. In particular, while the present invention has been described as being implemented in a network comprising one or more MAN switches, such as edge switch 1 **115**, edge switch 2 **130**, core MAN switches in VMAN CORE **120**, and customer domains such as the 802.1Q domains **110/130**, some of the logic may be distributed in other components of a network or internetwork application.

For example, embodiments of the invention may be represented as a software product stored on a machine-accessible medium (also referred to as a computer-readable medium or a processor-readable medium). The machine-accessible medium may be any type of magnetic, optical, or electrical storage medium including a diskette, CD-ROM, memory device (volatile or non-volatile), or similar storage mechanism. The machine-accessible medium may contain various sets of instructions, code sequences, configuration information, or other data. As an example, the procedures described herein for a encapsulating a data packet by edge switch 1 **115**, or forwarding a VMAN tagged frame by a core switch in VMAN CORE **120** can be stored on the machine-accessible medium. Those of ordinary skill in the art will appreciate that other instructions and operations necessary to implement the described invention may also be stored on the machine-accessible medium.

Thus, the present invention is not limited by the details described. Instead, the present invention can be practiced with modifications and alterations within the spirit and scope of the appended claims.
